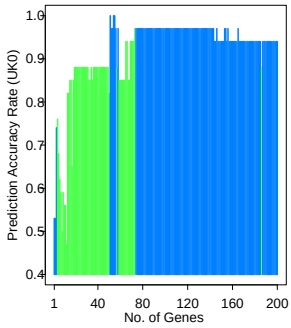
LEGENDS FOR SUPPLEMENTARY FIGURES

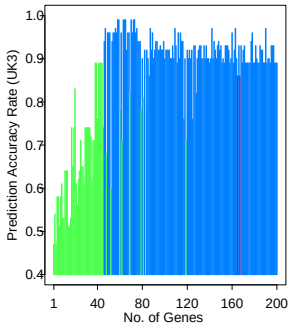
Figure S1 (page 2): Prediction accuracy rates for Leukemia Variants UK0 - UK3, using preprocessing method (A). The 8 Figures on the left are based on BIC model selection, while the right half are based on AIC model selection. For these figures as well as those in S2 and S3, red, blue and green lines indicate that the number of classes selected by BIC or AIC is $M < 3$, $M = 3$, and $M > 3$, respectively.

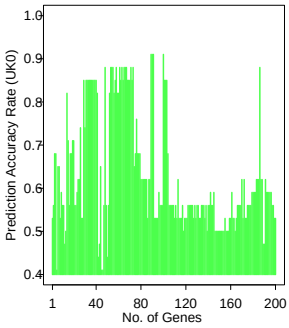
Figure S2 (page 3): Prediction accuracy rates for Leukemia variants UK0 - UK1, using preprocessing method (B). The 4 figures on the left half are based on BIC model selection, while the 4 figures on the right half are based on AIC model selection.

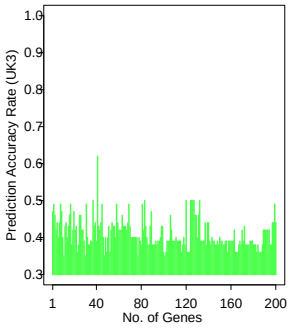
Figure S3 (page 4): Prediction accuracy rates for Leukemia variants UK0 - UK1, using preprocessing method (C). The 4 figures on the left half are based on BIC model selection, while the 4 figures on the right half are based on AIC model selection.

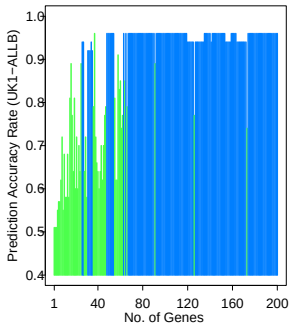
Figure S4 : Results for simulation studies based on Leukemia variants UK0 - UK3, using preprocessing method (A). The 8 figures on columns 1 and 3 present percents of datasets for which BIC and AIC select model with M components. Red, blue, green lines indicate datasets where BIC selects models with $M < 3$, $M = 3$, and $M > 3$ components, respectively, while black lines indicate datasets where AIC selects model with M = 3 components. The 8 figures on columns 2, 4 present several summary statistics of the number of errors among the datasets for which M=3 is correctly inferred using BIC. Green, red, and blue curves indicate the first quartile, median, and the third quartile, respectively, of the errors. Note that the gaps in the summary statistics plots for UK2-AML-ALLB and UK2-ALLB-ALLT are due to the fact that there is no datasets whose number of classes is correctly inferred.

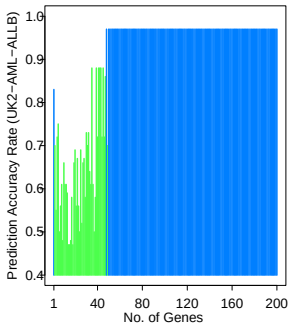

Prediction Accuracy Rate (UK0)
No. of Genes

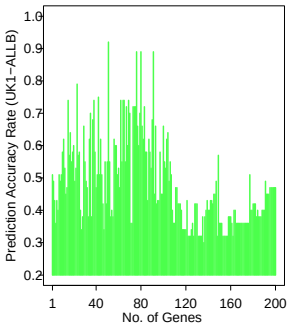

Prediction Accuracy Rate (UK3)
No. of Genes

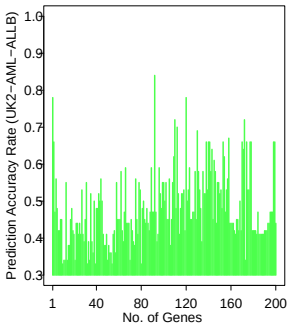

Prediction Accuracy Rate (UK0)
No. of Genes

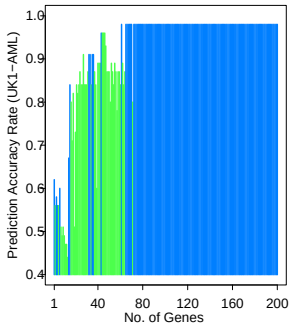

Prediction Accuracy Rate (UK3)
No. of Genes

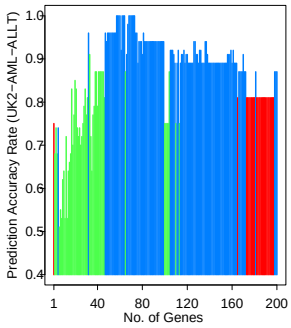

Prediction Accuracy Rate (UK1−ALLB)
No. of Genes

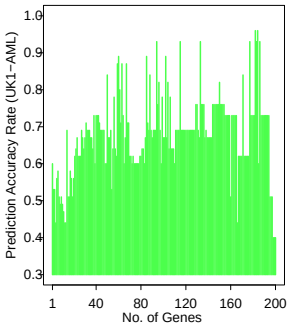

Prediction Accuracy Rate (UK2−AML−ALLB)
No. of Genes

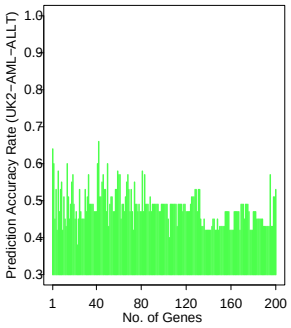

Prediction Accuracy Rate (UK1−ALLB)
No. of Genes

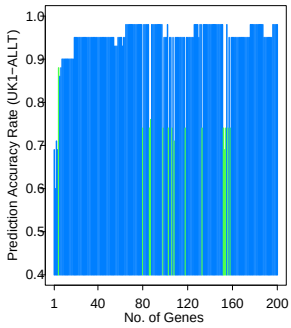

Prediction Accuracy Rate (UK2−AML−ALLB)
No. of Genes

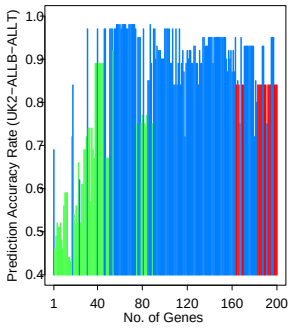

Prediction Accuracy Rate (UK1−AML)
No. of Genes

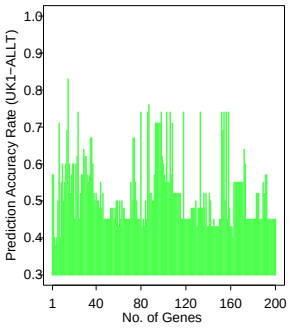

Prediction Accuracy Rate (UK2−AML−ALLT)
No. of Genes

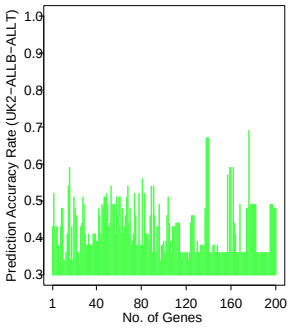

Prediction Accuracy Rate (UK1−AML)
No. of Genes

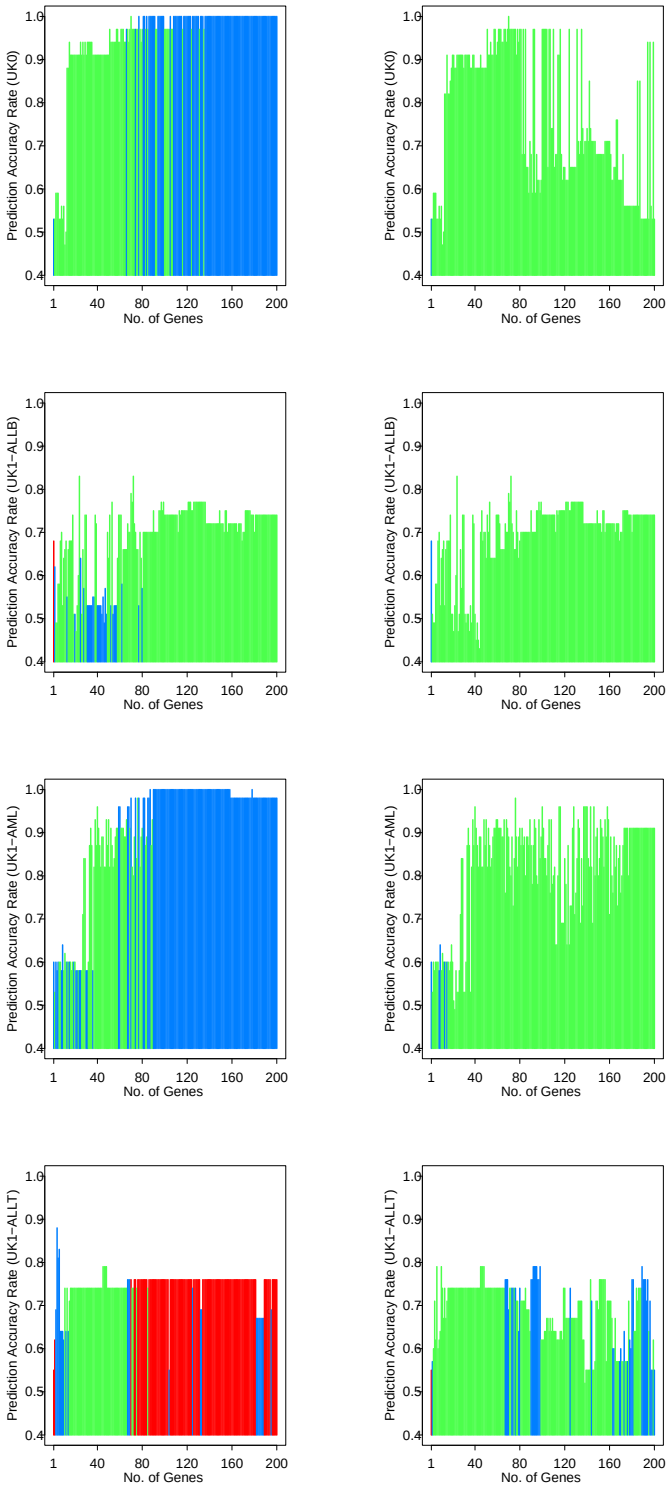

Prediction Accuracy Rate (UK2−AML−ALLT)
No. of Genes

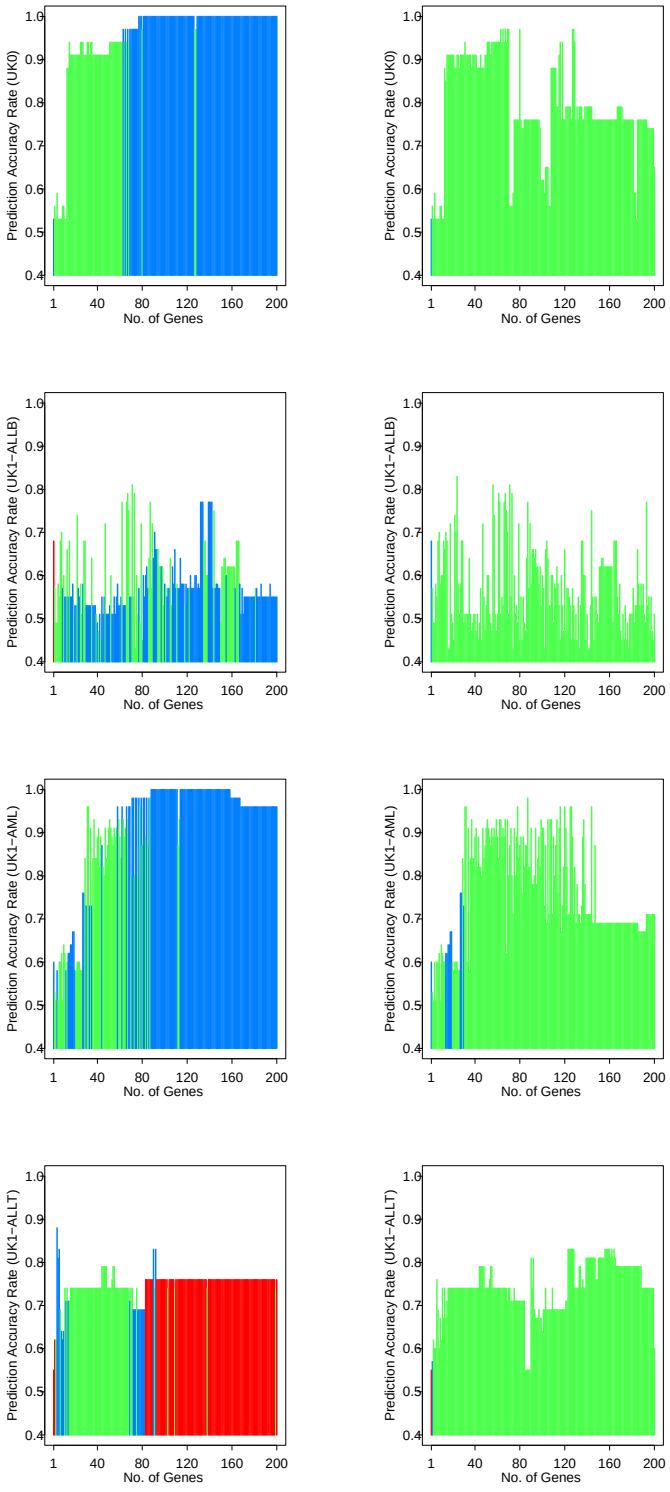

Prediction Accuracy Rate (UK1−ALLT)
No. of Genes

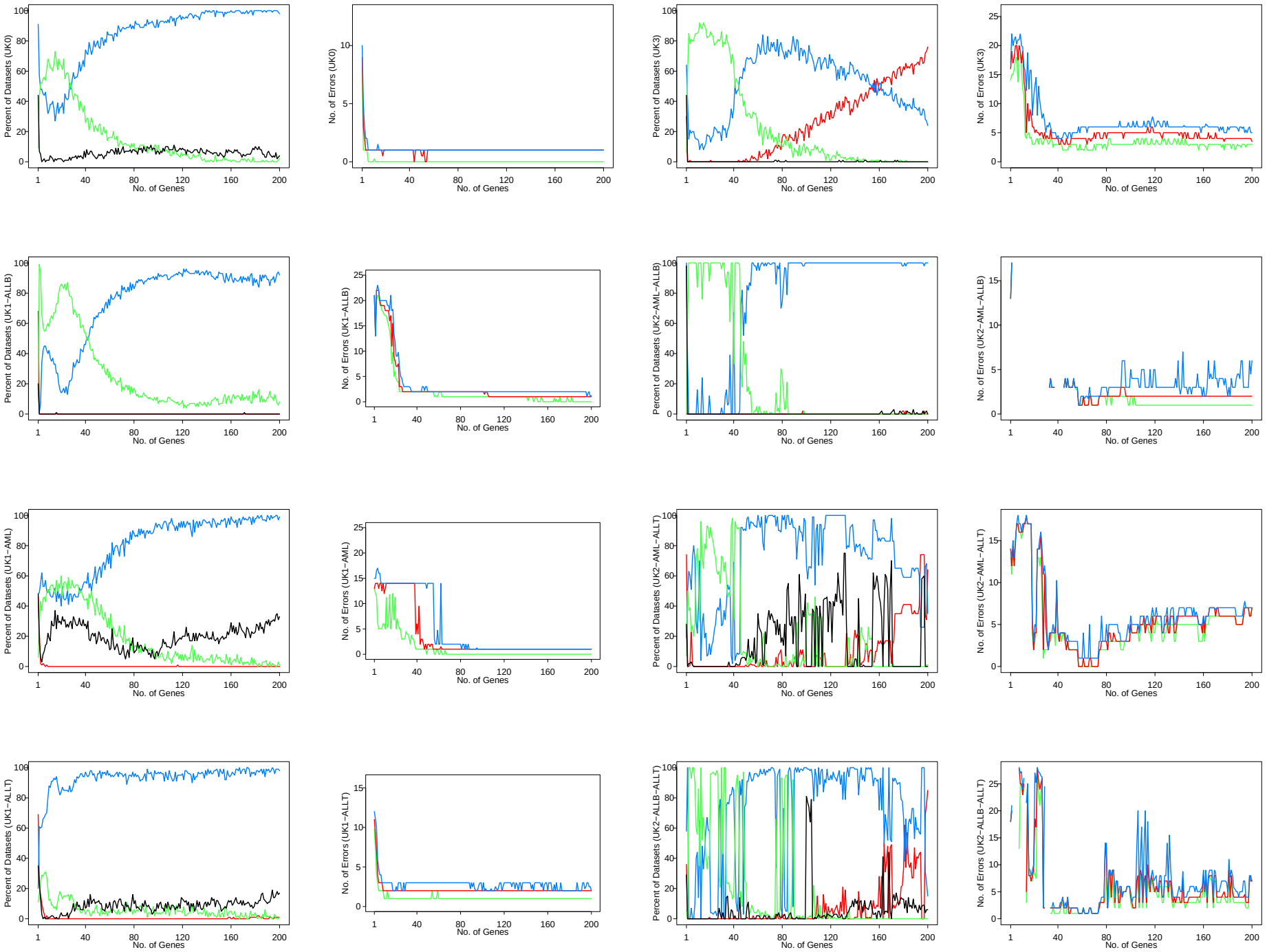

Prediction Accuracy Rate (UK2−ALLB−ALLT)
No. of Genes


Prediction Accuracy Rate (UK1−ALLT)
No. of Genes


Prediction Accuracy Rate (UK2−ALLB−ALLT)
No. of Genes


Prediction Accuracy Rate (UK0)
Prediction Accuracy Rate (UK1−ALLB)
Prediction Accuracy Rate (UK1−AML)
Prediction Accuracy Rate (UK1−ALLT)
No. of Genes
1.0
0.9
0.8
0.7
0.6
0.5
0.4
1
40
80
120
160
200


Prediction Accuracy Rate (UK0)
Prediction Accuracy Rate (UK1−ALLB)
Prediction Accuracy Rate (UK1−AML)
Prediction Accuracy Rate (UK1−ALLT)
No. of Genes
1.0
0.9
0.8
0.7
0.6
0.5
0.4
1
40
80
120
160
200


Percent of Datasets (UK0)
No. of Errors (UK0)
Percent of Datasets (UK3)
No. of Errors (UK3)
Percent of Datasets (UK1−ALLB)
No. of Errors (UK1−ALLB)
Percent of Datasets (UK2−AML−ALLB)
No. of Errors (UK2−AML−ALLB)
Percent of Datasets (UK1−AML)
No. of Errors (UK1−AML)
Percent of Datasets (UK2−AML−ALLT)
No. of Errors (UK2−AML−ALLT)
Percent of Datasets (UK1−ALLT)
No. of Errors (UK1−ALLT)
Percent of Datasets (UK2−ALLB−ALLT)
No. of Errors (UK2−ALLB−ALLT)
No. of Genes

C